

# Hashing the message with cells

Jongmin Kim &amp; Elisa Franco

 Check for updates

The programming of multicellular processes for biological computation requires increasingly complex genetic circuit design. Through automated circuit design, it is now possible to systematically break down complex response functions into dozens of interconnected simplified logic circuits, each embeddable in distinct cellular strains.

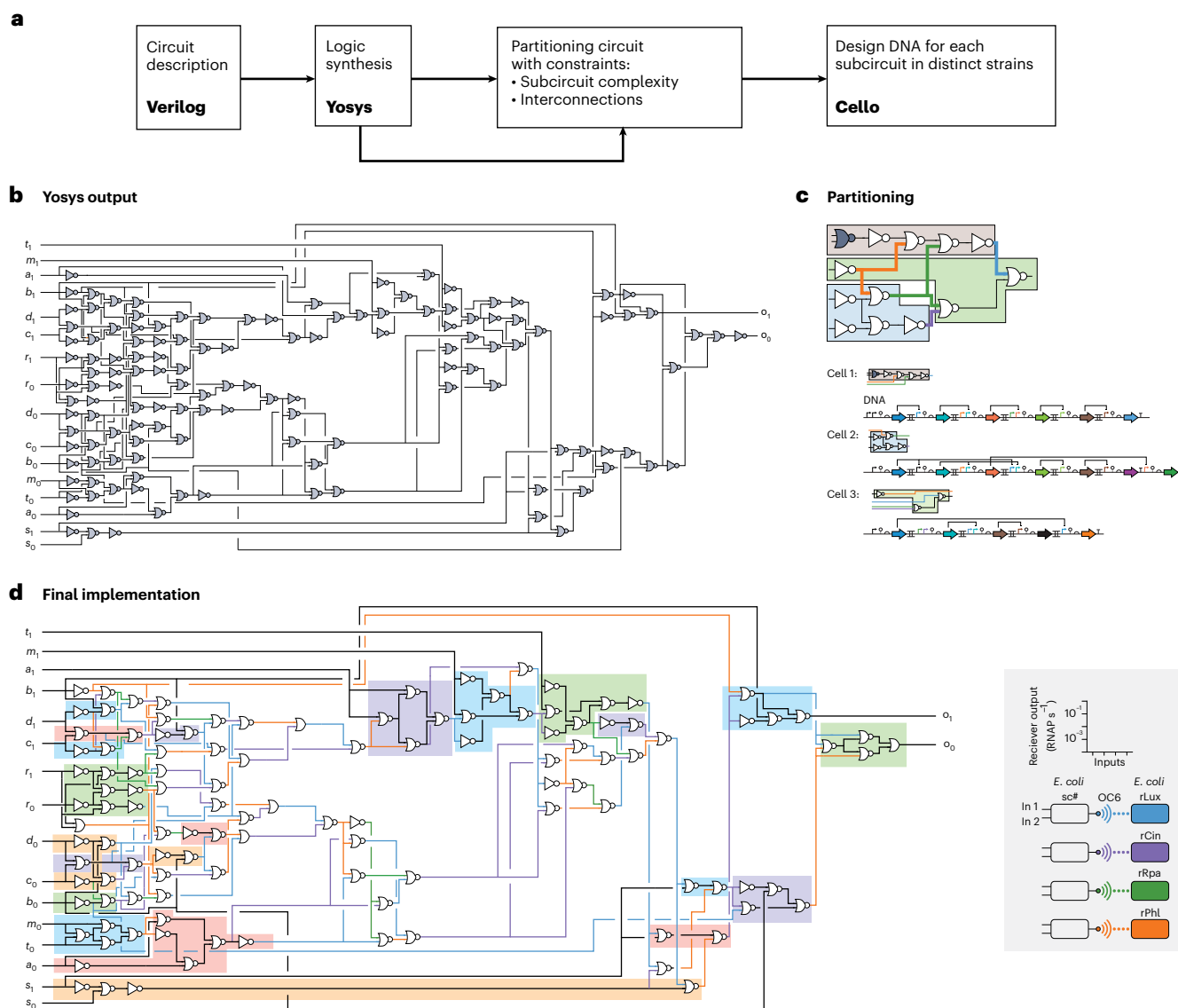
Synthetic biology aims to take the rational design principles of engineering and apply them to program a cell to execute a desired computational operation<sup>1</sup>. This strategy has resulted in the construction of increasingly complex genetic circuits, such as digital logic gates, memory devices, analog computation and dynamic circuits<sup>2</sup>. These have been further integrated into biotechnological applications such as metabolic pathway engineering and living biotherapeutics. Yet, the expression of regulatory components burdens the cell, leading to growth defects, circuit failures and evolutionary breakage<sup>3</sup>. Furthermore, synthetic biological circuit design has relied for some time on manual curation through lengthy processes of trial and error. These issues underscore the need for new strategies to scale up the complexity of synthetic genetic circuits.

In recent years, the paradigm for circuit design has shifted to computer-aided automated designs. An end-to-end design framework for logic circuit construction from the circuit specification to DNA sequences can provide an essential component for scaling up synthetic circuit designs (Fig. 1a). For example, an automated design software, Cello, builds on electronic design automation principles, allowing specification by Verilog, a hardware description language used to describe circuit behavior, and mapping with a predefined parts library<sup>4</sup>. Cello was previously used to design 60 circuits for *Escherichia coli*, and 45 circuits (up to 10 regulators and 55 parts) performed correctly in the assays<sup>4</sup>. The hope was that this approach would be more cost-effective and would boost synthetic biological applications by removing many errors and sub-optimizations that would have required manual tuning. Although this approach has been useful, it carried certain limitations such as allowing a single gate type (NOR gate) with a fixed architecture and lacking libraries for organisms other than *E. coli*. A later version addressed these limitations to provide more flexibility in the choice of gate architecture, design rules and chassis organisms<sup>5</sup>.

Now, in this issue of *Nature Chemical Biology*, Padmakumar et al.<sup>6</sup> address both automation strategies for complex genetic circuit design and the burden of such circuits on cells. They report the development of an automated genetic circuit design framework that combines Verilog for logic circuit design, Yosys for circuit synthesis and a new version of Cello for the implementation of circuits using genetic components. This framework makes it possible to scale up the complexity and stability of genetic circuits by implementing parallel processing of multiple

programs distributed among a community of cells<sup>6</sup>. The team used the automated circuit design framework to distribute computing tasks for the MD5 hashing algorithm, an early predecessor for cryptographic functions (Fig. 1b). The design of this large circuit presented some unique challenges including the partitioning of circuits to different cells to limit the maximum number of gates introduced into individual cells, a strategy that reduces metabolic burden due to exogenous gene expression. Similarly, the number of orthogonal communication channels for cell–cell communication must be limited to the four available molecular signals. The authors turned to graph partitioning and node-coloring algorithms to tackle this challenge, subdividing a large gene network into an equivalent interconnection of simpler networks each embedded in distinct cell lines (Fig. 1c). Together, one iteration of the hashing algorithm required 110 logic gates, which were partitioned across 66 strains of *E. coli*, requiring the introduction of 1.1 Mb of recombinant DNA into the genomes. Of these strains, some were identical, requiring the construction of 41 different strains. The largest subcircuit required 2 inputs, 8 gates and 3 communication outputs with a total number of 41 genes (24 regulatory) and 31 kb of subcircuit DNA. Finally, the strains that contain subcircuits were individually verified and the signal propagation of three layers was validated (Fig. 1d). Importantly, none of the cells containing the subcircuits exhibited a substantial growth defect, in contrast to earlier works<sup>4,7</sup>. Two main factors may contribute to this enhanced stability of synthetic circuits. First, the circuits were incorporated into the genome of cells, bypassing the danger of plasmid loss that leads to circuit breakage<sup>8</sup>. Second, a new library of phage repressors with known promoters and unique operator sequences was used for circuit construction, which appears to be less toxic and variable than previous TetR family repressors<sup>9</sup>. The identification of stable landing pads in the genome and a large potential library of phage repressors beyond the 12 used in this study point to a potentially general strategy to improve synthetic circuit stability moving forward.

Despite the demonstration of the design pipeline for the largest and most complex synthetic genetic circuits, there are some challenges that remain to be addressed. First, the lag time to accumulate the cell–cell signaling molecules and to induce response in receiver cells is estimated to be 84 hours for a single iteration of the algorithm, requiring approximately 200 days for 64 iterations required to complete the hashing of the input message by the MD5 algorithm. Second, as the number of quorum signals is limited, the re-use of cell–cell communication channels limits the experimental verification of multiple layers. Third, without synchronization strategies, co-culture of strains with different subcircuits can lead to faults when the signal skips layers. Alternative and mitigating strategies such as gate designs with larger fan-in and fan-outs to reduce the circuit depth, delivery of programs by direct contacts to limit the diffusive signal loss and editing of genomes for next-generation-sequencing-compatible measurements for large-scale parallel data acquisition may provide complementary approaches. Finally, the work of Padmakumar et al.<sup>6</sup> not only improves our understanding of the design principles for synthetic genetic circuits across cells, but also provides strong support for automated



**Fig. 1 | A divide and conquer strategy to embed circuits performing complex computations in multicellular systems. a**, Workflow of circuit design, partitioning and implementation using distinct bacterial strains. **b**, Yosys circuit implementation of the hashing algorithm. **c**, The circuit in **b** is partitioned in simplified interconnected units that are each easily implementable in a particular

strain while minimizing the number of components required. **d**, Implementation of the initial circuit; colors indicate distinct strains used to implement each subcircuit (sc). RNAP, RNA polymerase. Panels **b–d** adapted from ref. 6, Springer Nature Ltd.

design software as a powerful approach in scaling up the complexity of synthetic circuits for future applications.

Jongmin Kim <sup>1</sup> & Elisa Franco <sup>2,3</sup>

<sup>1</sup>Department of Life Sciences, Pohang University of Science and Technology, Pohang, Republic of Korea. <sup>2</sup>Department of Mechanical and Aerospace Engineering, University of California at Los Angeles, Los Angeles, CA, USA. <sup>3</sup>Department of Bioengineering, University of California at Los Angeles, Los Angeles, CA, USA.

e-mail: [jongmin.kim@postech.ac.kr](mailto:jongmin.kim@postech.ac.kr); [efranco@seas.ucla.edu](mailto:efranco@seas.ucla.edu)

Published online: 20 January 2025

## References

- Liang, J., Luo, Y. & Zhao, H. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **3**, 7–20 (2011).
- Cubillos-Ruiz, A. et al. *Nat. Rev. Drug Discov.* **20**, 941–960 (2021).
- Nikolados, E.-M., Weiße, A. Y., Ceroni, F. & Oyarzún, D. A. *ACS Synth. Biol.* **8**, 1231–1240 (2019).
- Nielsen, A. A. K. et al. *Science* **352**, aac7341 (2016).
- Jones, T. S., Oliveira, S. M. D., Myers, C. J., Voigt, C. A. & Densmore, D. *Nat. Protoc.* **17**, 1097–1113 (2022).
- Padmakumar, J. P. et al. *Nat. Chem. Biol.* <https://doi.org/10.1038/s41589-024-01730-1> (2024).
- Shin, J., Zhang, S., Der, B. S., Nielsen, A. A. & Voigt, C. A. *Mol. Syst. Biol.* **16**, e9401 (2020).
- Shao, B. et al. *Nat. Commun.* **12**, 1475 (2021).
- Lemire, S., Yehli, K. M. & Lu, T. K. *Annu. Rev. Virol.* **5**, 453–476 (2018).

## Competing interests

The authors declare no competing interests.